6
Hansenula polymorpha

Hyun Ah Kang and Gerd Gellissen

List of Genes

Gene	Encoded gene product
CAT	catalase
CNE1	calnexin
CTT1	catalase T (*S. cerevisiae*)
CMK2	calmodulin-dependent kinase
CWP1	Cell wall mannoprotein
DAK1	dihydroxyacetone kinase
DAS	dihydroxyacetone synthase
FLD1	formaldehyde dehydrogenase
FMD	formate dehydrogenase
GAM1	glucoamylase (*Schwanniomyces occidentalis*)
GAP	glyceraldehyde-3-phosphate dehydrogenase
GAS1	GPI-anchored surface glycoprotein
LEU2	ß-isopropyl malate dehydrogenase
MFα1	α mating factor (*S. cerevisiae*)
MOX	methanol oxidase
PHO1	acid phosphatase
PMA1	plasma membrane ATPase
PMR1	P-type Ca^{2+} transport ATPase
SED1	Cell surface glycoprotein (Suppressor of Erd2 Deletion)
TIP1	Cell wall mannoprotein (Temperature shock-inducible protein)
TPS1	trehalose-6-phosphate synthase
TRP3	indole-3-glycerol-phosphate synthase
URA3 (*ODC1*)	ornithidine decarboxylase
YNT1	nitrate transporter
YNI1	nitrite reductase
YNR1	nitrate reductase

Production of Recombinant Proteins. Novel Microbial and Eukaryotic Expression Systems. Edited by Gerd Gellissen

ISBN: 3-527-31036-3

6.1
History, Phylogenetic Position, Basic Genetics and Biochemistry of *H. polymorpha*

A limited number of yeast species are able to utilize methanol as a sole energy and carbon source. The range of methylotrophic yeasts includes *Candida boidinii*, *Pichia methanolica*, *Pichia pastoris* (see Chapter 7) and *Hansenula polymorpha* (Gellissen 2000). The latter two are distinguished by a growing track record of application in heterologous gene expression. In particular, *H. polymorpha* has found successful application in the industrial production of heterologous proteins, as detailed in later sections of this chapter (Gellissen 2002, Goengerich et al. 2004). Since *H. polymorpha* is the more thermotolerant of the two yeasts, it might also be better suited as source and for the production of proteins considered for crystallographic studies. In basic research it is used as model organism for research in peroxisomal function and biogenesis, as well as nitrate assimilation (Gellissen and Veenhuis 2001; van der Klei and Veenhuis 2002; Swerio 2002; Gellissen 2002). Again, the presence of a nitrate assimilation pathway is a feature not shared by *P. pastoris*.

The first methylotrophic yeast described was *Kloeckera* sp. No 2201, later re-identified as *Candida boidinii* (Ogata et al. 1969). Subsequently, other species, including *H. polymorpha*, were identified as having methanol-assimilating capabilities (Hazeu et al. 1972). Three basic strains of this species with unclear relationships, different features, and independent origins are used in basic research and biotechnological applications: strain CBS4732 (CCY38-22-2; ATCC34438, NRRL-Y-5445) was initially isolated by Morais and Maia (1959) from soil irrigated with waste water from a distillery in Pernambuco, Brazil. Strain DL-1 (NRRL-Y-7560; ATCC26012) was isolated from soil by Levine and Cooney (1973). The strain named NCYC495 (CBS1976; ATAA14754, NRLL-Y-1798) is identical to a strain first isolated by Wickerham (1951) from spoiled concentrated orange juice in Florida and initially designated *Hansenula angusta*. Strains CBS4732 and NCYY495 can be mated, whereas strain DL-1 cannot be mated with the other two (K. Lahtchev, personal communication).

The genus *Hansenula* H. et P. Sydow includes ascosporogenic yeast species exhibiting spherical, spheroidal, ellipsoidal, oblong, cylindrical, or elongated cells. One to four ascospores are formed. Ascigenic cells are diploid arising from conjugation of haploid cells. The genus is predominantly heterothallic. *H. polymorpha* is probably homothallic, exhibiting an easy interconversion between the haploid and diploid states (Tennasson et al. 1960; Middelhoven 2002). Since the morphological characteristics of the *Hansenula* species are shared by species of the genus *Pichia* Hansen, Kurtzman (1984), after performing DNA/DNA reassociation studies, proposed to merge both genera and transfer *Hansenula* species with hat-shaped ascospores to *Pichia* Hansen emend Kurtzman, although *Hansenula* spp. can grow on nitrate and *Pichia* spp. cannot. Kurtzman and Robnett (1998) provided a phylogenetic tree in which nitrate-positive and nitrate-negative *Pichia* are clustered, demonstrating the unreliability of nitrate assimilation for prediction of kinship. The leading taxonomy monographs follow this proposal, re-naming *H. polymorpha* as *Pichia angusta* (Kurtzman and Fell 1998; Barnett et al. 2000). However, the merging of the genera is still criticized by some taxonomists, and there are arguments for mainiain-

Fungi (Kingdom)
Eumycota (Division)
Ascomycotina (Subdivision)
(Hemiascomycetes (Class)
Endomycetales (Order)
Saccharomycetaceae (Family)
(Saccharomycetoideae (Sub-family)
Hansenula (Genus)
Hansenula polymorpha (Species)

Fig. 6.1 Taxonomy of *Hansenula polymorpha* (after Kreger-van Rij 1984)

ing the established and popular name *Hansenula polymorpha* (Middelhoven 2002; Sudbery 2003).

Among a wealth of biochemical and physiological characteristics, some selected features are presented in the following sections; for a more comprehensive view, the reader is referred to the chapters of a recent monograph (Gellissen 2002). Some strains of *H. polymorpha* can tolerate temperatures of 49 °C and higher (Teunisson et al. 1960; Reinders et al. 1999). It was shown that trehalose synthesis is not required for growth at elevated temperatures, but that it is necessary for normal acquisition of thermotolerance (Reinders et al. 1999). The thermoprotective compound trehalose accumulates in large amounts in cells grown at high temperatures. The synthetic steps for trehalose synthesis have been detailed for *H. polymorpha*. The *TPS1* gene encoding trehalose 6-phosphate synthase is the key enzyme gene of this pathway (Romano 1998; Reinders et al. 1999). Transcripts of this gene were found to be present in high quantities in cells grown at normal temperatures, and to be especially abundant when grown at elevated temperatures (Reinders et al. 1999). Accordingly, the *TPS1*-derived promoter provides an attractive element to drive constitutive heterologous gene expression which can be further boosted at temperatures above 42 °C (Ammel et al. 2000; Suckow and Gellissen 2002; see also the following sections).

The capability of *H. polymorpha* to grow on methanol as a sole energy and carbon source is enabled by a methanol utilization pathway that is shared by all known methylotrophic yeasts (Tani 1984; Yurimoto et al. 2002; see also Chapter 7 on *P. pastoris*). Growth on methanol is accompanied by a massive proliferation of peroxisomes in which the initial enzymatic steps of this pathway take place (Figure 6.2) (Gellissen and Veenhuis 2001; van der Klei and Veenhuis 2002; Yurimoto et al. 2002).

The enzymatic steps of the compartmentalized methanol metabolism pathway are detailed in Figure 6.3. For more comprehensive information, the reader is referred to Yurimoto et al. (2002).

During growth on methanol, key enzymes of the methanol metabolism are present in high amounts. An especially high abundance can be observed for methanol

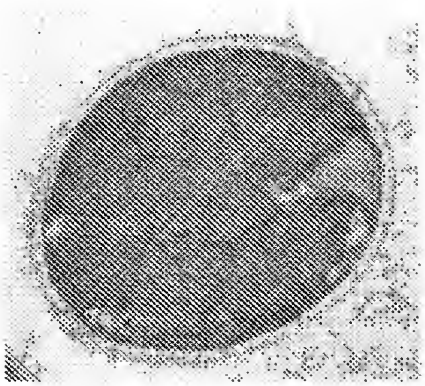


Fig. 6.2 *Hansenula polymorpha* cell. The cells grown in a methanol-limited chemostat are crowded with peroxisomes (Courtesy of M. Veenhuis).

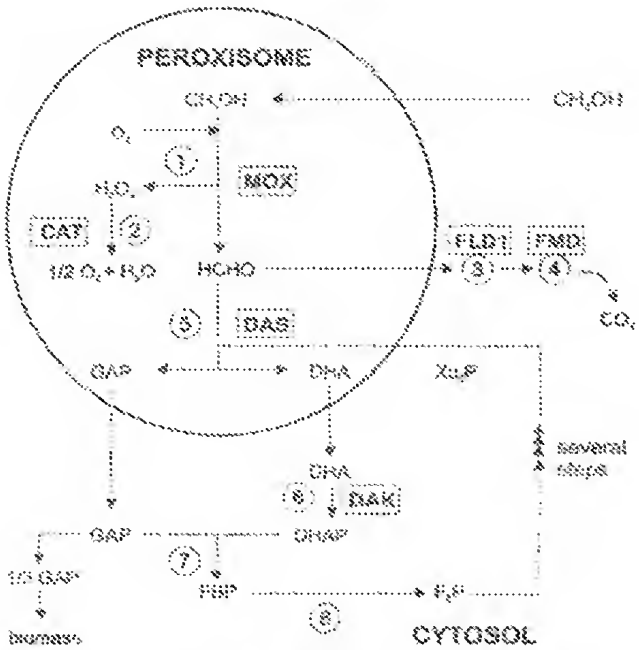


Fig. 6.3 The methanol utilization pathway and its compartmentalization in methylotrophic yeasts. (Modified after Gellissen 2000; Yurimoto et al. 2002.) *1* Methanol is oxidized by alcohol oxidase to generate formaldehyde and hydrogen peroxide. *2* The toxic hydrogen peroxide is decomposed by catalase to water and oxygen. *3, 4* Within a dissimilatory pathway the formaldehyde is oxidized by two subsequent dehydrogenase reactions to carbon dioxide, catalyzed by a formaldehyde dehydrogenase (FLD) and a formate dehydrogenase (FMD or FDH). *5* For assimilation, the formaldehyde reacts with xylulose 5-phosphate (Xu_5P) by the action of dihydroxyacetone synthase (DHAS) to generate the C_3 compounds glyceraldehyde 3-phosphate (GAP) and dihydroxyacetone (DHA). *6* DHA is phosphorylated by dihydroxyacetone kinase (DHAK) to dihydroxyacetone phosphate (DHAP). *7* GAP and DHAP yield in an aldolase reaction fructose 1,6-biphosphate (FBP). *8* In further steps of the pentose phosphate cycle, fructose 5-phosphate and xylulose 5-phosphate are finally generated. Identified and characterized genes of the *H. polymorpha* methanol utilization pathway are boxed and shown in the pathway position of the encoded enzymes. The genes are *MOX* (Ledeboer et al. 1985), *DAS* (Janowicz et al. 1985), *CAT* (Didion and Roggenkamp 1992), *DAK* (Tikhomirova et al. 1988), *FLD1* (Baerends et al. 2002), and *FMD* (Hollenberg and Janowicz 1988).

oxidase (MOX), formiate dehydrogenase (FMD), and dihydroxyacetone synthase (DHAS) (Gellissen et al. 1992a). The presence of these enzymes is regulated at the transcriptional level of the respective genes. Gene expression is subject to a carbon source-dependent repression/derepression/induction mechanism conferred by inherent properties of their promoters. Promoters are repressed by glucose, derepressed by glycerol, and induced by methanol. Again, these promoter elements – and in particular the elements derived from the *MOX* and the *FMD* genes – constitute attractive components for the control of heterologous gene expression that can be regulated by carbon source addition (see forthcoming sections). The possibility of eliciting high promoter activity with glycerol as sole carbon source and even with limited addition of glucose (glucose starvation) is unique among the methylotrophic yeasts. In the related species *C. boidinii*, *P. methanolica*, and *P. pastoris*, the active status of the promoter is strictly dependent on the presence of methanol or methanol derivatives (Gellissen 2000). However, this does not seem to be an inherent promoter characteristic; rather, it rather depends on the cellular environment of the specific host as upon transfer into *H. polymorpha* the *P. pastoris*-derived *AOX1* promoter is active under glycerol conditions (Raschke et al 1996; Rodríguez et al 1996).

6.2
Characteristics of the *H. polymorpha* Genome

As pointed out before, there are several *H. polymorpha* strains available. In the following section we focus on strains CBS4732 (CCY38–22-2; ATCC34438; NRRL-Y-5445) and DL-1 (NRRL-Y-7560; ATCC26012) which are the two ancestor strains of the preferred *H. polymorpha* hosts employed in heterologous gene expression. Data on karyotyping are available for both strains (Figure 6.4; Table 6.1).

Pulsed-field gel electrophoresis of *H. polymorpha* chromosomes revealed that both strains have six chromosomes, ranging from 0.9 to 1.9 Mbp, but the electrophoretic patterns of their chromosomes were quite different. The scientific and industrial significance of strain CBS4732 is now met by the recent characterization of its entire

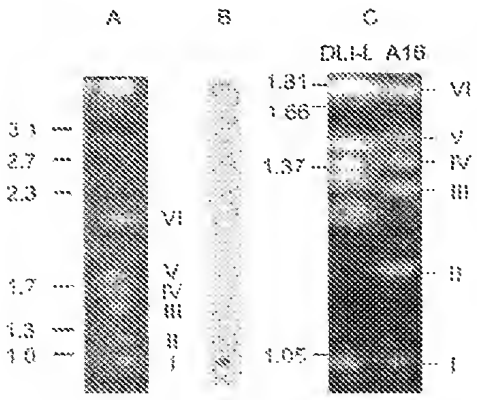


Fig. 6.4 Electrophoretic karyotype of *H. polymorpha* strains CBS4732 and DL-1. A) Chromosome pattern of *H. polymorpha* strain CBS4732 separated by pulsed-field gel electrophoresis (Schwarz and Cantor 1984) using the Pulsaphor apparatus (Pharmacia). Details of the separation conditions are provided elsewhere (Waschk et al. 2002).
B) Chromo Blot. The separated chromosomes were transferred to a nylon membrane and hybridized to a labeled *URA3* probe. A signal was obtained exclusively with chromosome I. (Modified and supplemented according to Waschk et al. 2002.)
C) Chromosome pattern of strain DL-1 (M) Ok, HA Kang, unpublished results). Details of the separation conditions will be published elsewhere.

Tab. 6.1 Chromosomal localization of several cloned genes in strains CBS4732.

Cloned gene/sequence used as specific probe	*Function*	*Reference*	*Chromosome no.*
URA3	Orotidine-5'-phosphate decarboxylase	Merckelbach et al. (1993)	I
CPY	Carboxypeptidase Y	Unpublished [1]	I
GAP	Glyceraldehyde 3-phosphate dehydrogenase	Unpublished [2]	I
rDNA (5.8S, 18S, 26S)	Ribosomal DNA	Klabunde [3]	II
HARS1	Autonomously replicating sequence 1	Ledeboer et al. (1986)	III
TPS1	Trehalose 6-phosphate synthase	Reinders et al. (1999)	IV
HLEU2	β-Isopropylmalate dehydrogenase	Agaphonov et al. (1994)	IV
MOX	Methanol oxidase	Ledeboer et al. (1985)	V
FMD	Formiate dehydrogenase	Hollenberg and Janowicz (1989)	VI

1) Bae JH, Kim HY, Sohn JH, Choi ES, Rhee SK. Accession number U67174.
2) Sohn JH, Choi ES, Rhee SK. Accession number U95625.
3) J Klabunde, personal communication.

genome (Ramezani-Rad et al. 2003). A few gene sequences are elucidated and can be compared for both strains (Table 6.2).

The sequence identity of the open reading frame (ORF) for the selected genes ranges between 94.5 and 97.2%, with an average value of 96.6%. The sequence differences are observed to be much more apparent at the 5'- and 3'-untranslated regions, which might be involved in controlling gene expression. This implies that

Tab. 6.2 Comparison of selected gene sequences from *H. polymorpha* strains CBS4732 (RB11) and DL-1.

Gene name	*Amino acid identity (%)*	*Nucleic acid identity (%)*	*Accession No. in GenBank**	*Reference*
CST13	96.7	95.8	AF454544	Kim et al. (2002)
CPY	98.0	95.9	U67174	KRIBB
GSH2	96.0	94.5	AF435121	KRIBB
MNN9	96.3	95.5	AF264786	Kim et al. (2001)
PMI40	97.9	94.9	AF454544	Kim et al. (2002)
PMR1	98.5	95.2	U92083	Kang et al. (1998)
YPT1	99.5	97.2	AF454544	Kim et al. (2002)
Average	97.6	96.6		

* The sequences of genes isolated from the DL-1 strain were obtained from GenBank and compared with those from the RB11 strain (Ramezani-Rad et al. 2003).

two strains are closely related, but have distinct genetic and physiological characteristics.

Several groups worldwide initiated studies on the CBS4732 genome several years ago. Included in the comparative genome analysis on 13 hemiascomycetous yeasts, part of the *H. polymorpha* genome sequence was established using a partial random sequencing strategy with a coverage of 0.3 genome equivalents. Using this approach, about 3 Mbp of sequencing raw data of the *H. polymorpha* genome was yielded (Feldmann et al. 2000). The recently terminated genome analysis aimed at a higher coverage using a BAC-to-BAC approach and resulting in the comprehensive genome analysis of this organism (Ramezani-Rad et al. 2003). For sequencing of *H. polymorpha* strain RB11, an *ade1*-derivative of strain CBS4732, a BAC library with approximately 17 × coverage was constructed in a pBACe3.6 vector according to Osoegawa et al. (1998, 1999). Details on base calling, assembly and editing are provided by Ramezani-Rad et al. (2003). Sequencing resulted in the characterization of 8.733 million base pairs assembled into 48 contigs. The derived sequence covers over 90% of the estimated total genome content of 9.5 Mbp located on six chromosomes which range in size from 0.9 to 2.2 Mbp (see Figure 6.4A). From the sequenced 8.73 Mb, a total of 5848 ORFs have been extracted for proteins longer than 80 amino acids (aa), and 389 ORFs smaller than 100 aa were identified. Likewise, 4771 ORFs have homologues to known proteins (81.6%). Calculation of the gene density and protein length, taking into account the gene numbers, showed an average gene density of one gene per 1.5 kb, and an average protein length of 440 amino acids. Ninety-one introns have been identified by homology to known proteins and confirmed by using GeneWise (Barney et al. 2000). Eighty tRNAs were identified, corresponding to all 20 amino acids. From approximately 50 rRNA clusters (Klabunde et al. 2002; Waschk et al. 2002; Klabunde et al. 2003), seven clusters have been fully sequenced. All clusters are completely identical and have a precise length of 5033 bp.

The main functional categories and their distribution in the gene set are manually predicted for energy, 4%; cellular communication, signal transduction mechanism, 3%; protein synthesis, 6%; cell rescue, defense and virulence, 4%; cellular transport and transport mechanisms, 9%; cell cycle and DNA processing, 9%; protein fate (folding, modification, destination), 17%; transcription, 13% and metabolism, 19%. A selection of the data obtained from the annotated sequence is provided in Tables 6.3 and 6.4.

Tab. 6.3 *Hansenula polymorpha* genome statistics.

Contigs:	48
Total length of contigs:	8 733 442 bp
Average contig length:	182 kb
No. of extracted ORFs:	5848
No. of ORFs <100 aa:	389
Average gene density:	1 gene/1.5 kb
Average gene size (start-stop):	1.3 kb (1320 nt)
Average protein length:	440 aa

Tab. 6.4 Functional categorization of genes.

Functional category	No. of ORFs	%
Metabolism	1114	19
Energy	231	4
Cell growth, cell division and DNA synthesis	518	9
Transcription	767	13
Protein synthesis	323	6
Protein destination	1014	17
Transport facilitation	423	7
Cellular transport and transport mechanisms	518	9
Control of cellular organization	417	7
Cellular communication/signal transduction	170	3
Cell rescue, defense, and virulence	260	4
Cell fate	282	5
Regulation of/interaction with cellular environment	184	3

6.3
N-linked glycosylation in *H. polymorpha*

The similarities between yeast and animal cell secretion pathways have made yeasts in general preferred microbial host systems for the production of human secretory proteins. A majority of human proteins with therapeutic potential are glycoproteins, and increasing evidence shows that oligosaccharides assembled on glycoproteins have profound effects on the properties of glycoproteins, such as antigenicity, specific activity, solubility, and stability. The initial processing of *N*-linked glycans on glycoproteins, which occurs in the endoplasmic reticulum (ER), is well conserved among eukaryotes and results in the core oligosaccharide, $Man_8GlcNAc_2$. However, further maturation of oligosaccharides in the Golgi apparatus is quite variable among organisms, even among yeast species (Gemmill and Trimble 1999). Yeasts elongate the core oligosaccharide mostly by addition of mannose, leading to the formation of core-sized structures ($Man_{>13}GlcNAc_2$) as well as hypermannose structures ($Man_{50-200}GlcNAc_2$) with extended poly-α-1,6 outer mannose chains, which are decorated with various carbohydrate side chains in a species-specific manner. In *Saccharomyces cerevisiae*, the linear backbone of the outer chain is often composed of 50 or more mannoses, highly branched by addition of α-1,2- linked mannoses and terminally capped with α-1,3-linked mannoses, generating heavily hypermannosylated glycoproteins. The outer chain also contains several mannosylphosphate residues, conferring the oligosaccharide with a negative net charge (Jigami and Odani 1999; Kim et al. 2004).

Compared to *S. cerevisiae*, the mannose outer chains of *N*-linked oligosaccharides generally appear to be much shorter in the methylotrophic yeasts *P. pastoris* and *H. polymorpha*, although extensive hyperglycosylation has also been reported in a few cases of recombinant glycoproteins produced in these yeasts (Scorer et al. 1993; Müller et al. 1998). Analyses of *N*-linked oligosaccharides added to native and recom-

binant glycoproteins from *P. pastoris* have indicated that the major oligosaccharide species in *P. pastoris* are $Man_{8-14}GlcNAc_2$ with short α-1,6 extensions. More significantly, *P. pastoris* oligosaccharides are reported to have no hyperimmunogenic terminal α-1,3 glycan linkages (Montesino et al. 1998; Bretthauer and Castellano 1999). Phosphomannose has been detected in both elongated and core oligosaccharides on some recombinant proteins of *P. pastoris* (Miele et al. 1997; Montesino et al. 1999). At present, very little information exists on the structural characteristics of *N*-linked oligosaccharides of *H. polymorpha*-derived glycoproteins. A comparative study on the glycosylation pattern of recombinant human α_1-antitrypsin produced in *S. cerevisiae*, *H. polymorpha*, and *P. pastoris* has suggested that the length of outer mannose chains attached to the recombinant protein in *H. polymorpha* was much shorter than in *S. cerevisiae*, but slightly longer than in *P. pastoris* (Kang et al. 1998b). A recent study on the structure of the oligosaccharides derived from the recombinant *Aspergillus niger* glucose oxidase (GOD) and the cell wall mannoproteins derived from *H. polymorpha* has revealed that most oligosaccharide species attached to the recombinant GOD have core-sized structures ($Man_{8-12}GlcNAc_2$) without terminal α-1,3-linked mannose residues (Kim et al. 2004). Therefore, the outer chain processing in the *N*-linked glycosylation pathway in *H. polymorpha* appears to be similar to that in *P. pastoris*, with the addition of short outer chains to the core and no terminal α-1,3-linked mannose addition (Figure 6.5).

In contrast to the yeast oligosaccharides composed solely of mannose, a variety of sugars including galactose, *N*-acetylgalactosamine and sialic acid, are added to oligosaccharides in mammals. The differences in glycan processing between yeasts and humans are a major limitation for yeasts to be used in the production of recombinant glycoproteins for therapeutic use. Glycoproteins derived from yeast expression systems contain nonhuman *N*-glycans of the high-mannose type, which are immunogenic in humans. Attempts have been made genetically to modify glycosylation processes in *S. cerevisiae* (Chiba et al. 1998) and *P. pastoris* (Callewaert et al. 2001), in order to trim the yeast *N*-glycans of the high-mannose type to the human glycans of the ($Man_5GlcNAc_2$) intermediate type. A more advanced achievement has been recently made genetically to re-engineer the glycosylation pathway of *P. pastoris* to produce the complex human *N*-glycan *N*-acetylglucosamine$_2$-mannose$_3$-*N*-acetylglucosamine$_2$ ($GlcNAc_2Man_3GlcNAc_2$) (Hamilton et al. 2003; see Chapter 7). Potentially, the development of *H. polymorpha* expression systems for proper glycosylation can be achieved as further understanding is gained of *H. polymorpha*-specific carbohydrate structure and processing sugar transferases. To date, only a few *H. polymorpha* genes and mutants related to protein glycosylation have been reported (Kang et al. 1998a; Agaphonov et al. 2001; Kim et al. 2001, 2002). Information from the *H. polymorpha* genome sequence (Blandin et al. 2000; Ramezani-Rad et al. 2003) will expedite the identification of genes that are predicted to be involved in the biosynthesis of sugar chains. The functional characterization of these genes should facilitate delineation of the *H. polymorpha*-specific *N*-glycosylation pathway, and this would provide valuable information for the development of glyco-engineering strategies in *H. polymorpha* to achieve the optimal glycosylation of recombinant proteins.

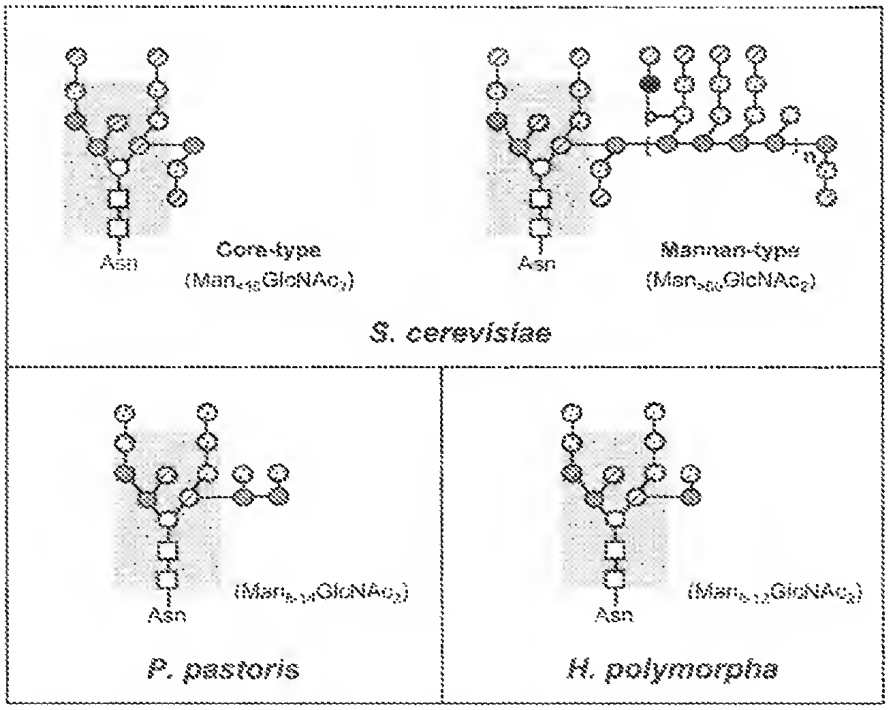


Fig. 6.5 The representative *N*-linked oligosaccharide structures assembled on the *S. cerevisiae*, *P. pastoris*, or *H. polymorpha*-derived glycoproteins. Information on the *S. cerevisiae* and *P. pastoris* oligosaccharides is from Gemmill and Trimble (1999).

6.4 The *H. polymorpha*-based Expression Platform

6.4.1 Transformation

Recombinant *H. polymorpha* strain generation requires special tools. Application of the commonly used *S. cerevisiae* 2 μm sequence with its predominantly episomal fate is restricted to a limited set of host strains (Gellissen and Hollenberg 1997). A variety of gene replacement approaches have been described, but these require DNA fragments with termini composing much longer target gene overlaps than those used in *S. cerevisiae* (Gonzalez et al. 1999; van Dijk et al. 2001). Typically, these overlaps must be in the magnitude of hundreds to thousands of base pairs. In addition, since the background caused by nonhomologous recombination is high, screening of more transformants than in *S. cerevisiae*-based systems is necessary to identify a strain with the desired integration/replacement (Kang et al. 2002). Plasmids harboring one of a set of several cloned sub-telomeric *ARS* sequences derived from the DL-1 strain have been described, which homologously integrate into a

genomic counterpart, resulting in the recombinant strains harboring single or multiple tandemly repeated copies at the respective sub-telomeric genomic locus (Sohn et al. 1999a; Kim et al. 2003). A set of vectors has been described to target the heterologous DNA to the rDNA locus of *H. polymorpha* (Klabunde et al. 2002, 2003; Suckow and Gellissen 2002; see also Chapter 13). Most commonly, plasmids harboring *HARS1* (*Hansenula ARS1*) as a replication signal are used to generate recombinant *H. polymorpha* strains. The fate of these plasmids in *H. polymorpha* has been extensively analyzed for both RB11 (derived from CBS4732; see the following section) and DL-1 strains (Suckow and Gellissen 2002; Kang et al. 2002). The strain generation by use of *HARS1* plasmids will therefore be described in more detail.

After transformation, cells are plated on selective media according to the selection marker gene present on the *HARS1* plasmid. Macroscopic colonies typically appear after 4–5 days of incubation at 37 °C. In this early phase, all cells of a colony harbor the plasmid as an episome at a low copy number; plasmid loss can be induced by cultivation on a nonselective medium. Colonies are then transferred to liquid selective medium and incubated under vigorous shaking for 24–48 h at 37 °C. This procedure is called the "passaging step". An aliquot of the dense culture is then used to inoculate fresh selective medium, and the incubation is repeated. After three to eight subsequently applied passaging steps, cells grow with the initially episomal plasmid integrated into the genome. In a particular single strain developed for the production of a hepatitis B vaccine, a recombination within the *FMD* locus was observed (U Dahlems, unpublished results). In order to separate these cells from those still harboring the plasmid as an episome, one or two "stabilization steps" must be performed in nonselective liquid media. If aliquots of these cultures are plated onto selective media, the resulting single colonies will exclusively represent strains harboring one to multiple tandemly repeated copies of the *HARS1* plasmid. The various individual strains can vary significantly in the expression rates of the foreign gene, present on the plasmid. However, typically only a few strains display very high levels of target protein. For a high probability of obtaining a "high producer" in the first approach, parallel generation of up to 150 strains is recommended. Once a suitable strain is identified, a so-called "supertransformation" can be performed using a second *HARS1* plasmid with a different selection marker gene. This second plasmid may contain either the same heterologous expression cassette as the first plasmid, or a different one. In the first case, strains would result which might display higher production rates of the target protein than the basic strain; in the second case, strains would be obtained co-expressing two different heterologous genes at variable but individually fixed relative expression rates (Gellissen 2000).

To summarize this procedure, the generation of recombinant strains in *H. polymorpha* is clearly more laborious than in other yeasts. However, these additional difficulties are balanced by two positive features which are highly favorable in biotechnological applications. First, heterologous gene expression in *H. polymorpha* can be controlled by homologous promoters of extraordinary strength. While the carbon source-regulated *MOX* and *FMD* promoters are derived from genes of the methanol degradation pathway, the *TPS1* promoter, derived from the trehalose 6-phosphate synthase gene of *H. polymorpha*, is constitutive with regard to different carbon

sources and can be influenced by different temperatures (Amuel et al. 2000). In combination with the obtained high copy numbers of the integrated *HARS1* plasmid (up to 100 copies per haploid genome may result from supertransformation), these strong promoters can provide very high expression rates of the heterologous gene in selected strains; indeed, for secreted phytase, product levels of up to 13.5 g L^{-1} have been obtained (Mayer et al. 1999). The second favorable feature of recombinant *H. polymorpha* is the unambiguous mitotic stability of the individual strains with regard to the copy numbers of the *HARS1* plasmids integrated, even upon cultivation on nonselective media for a long period of time. This stability has been well documented in several examples over periods of 800 generations. For plasmid examples, see Section 6.4.3 and Figures 6.7 and 6.8; for some product and process examples, see Section 6.5.

6.4.2 Strains

Starting with the three *H. polymorpha* parental strains, NCYC495, CBS4732, and DL-1 (see Section 6.1), some forty to fifty other strains have been derived. The DL-1 strain is not employed in classical genetic analyses, and no data are available regarding its ability to mate and to sporulate (Lahtchev 2002). The DL-1 strain has certain advantages in that it has a higher growth rate and adapts more quickly to culture media than the other parental strains; additionally, DL-1 strains have a higher frequency of homologous recombination than other strains (Kang et al. 2002; Lahtchev 2002). The inability of the DL-1 strain to copulate makes this strain inconvenient for classical genetic manipulation exploiting meiotic segregation. However, the relatively high frequency of homologous recombination in the DL-1 strain enables the application of several molecular genetic techniques developed in *S. cerevisiae* to be used in the organism. Several host strains suitable for heterologous protein expression, including auxotrophic mutants, protease-deficient strains, and mox-negative strains, have been constructed in the DL-1 strain, mostly using gene disruption techniques (see the list of strains in Appendix A6.1). A pop-out cassette using *HpURA3* as a selection marker has been constructed to recover the auxotrophic marker for the subsequent gene disruption or for subsequent transformation with expression vectors (Kang et al. 2002). Combined with fusion PCR and in-vivo recombination, the use of the *HpURA3* pop-out cassette becomes more simplified in constructing null mutant strains with disruption of the gene of interest (see Figure 6.6). This approach eliminates the time-consuming steps of ligation and sub-cloning, which are otherwise required for the construction of a gene deletion cassette.

Most classical genetic techniques have been performed using NCYC495, which shows both mating and sporulation (Lahtchev 2002). Unlike the other two parental strains, NCYC495 does not grow well on methanol-containing media and therefore does not have the strong methanol-induced promoters available to the other strains for gene expression. Instead, NCYC495 has other interesting applications, including its employment for the study of nitrate assimilation mentioned previously (Siverio 2002). Cells from CBS4732 grow well on methanol, and show strong mating and

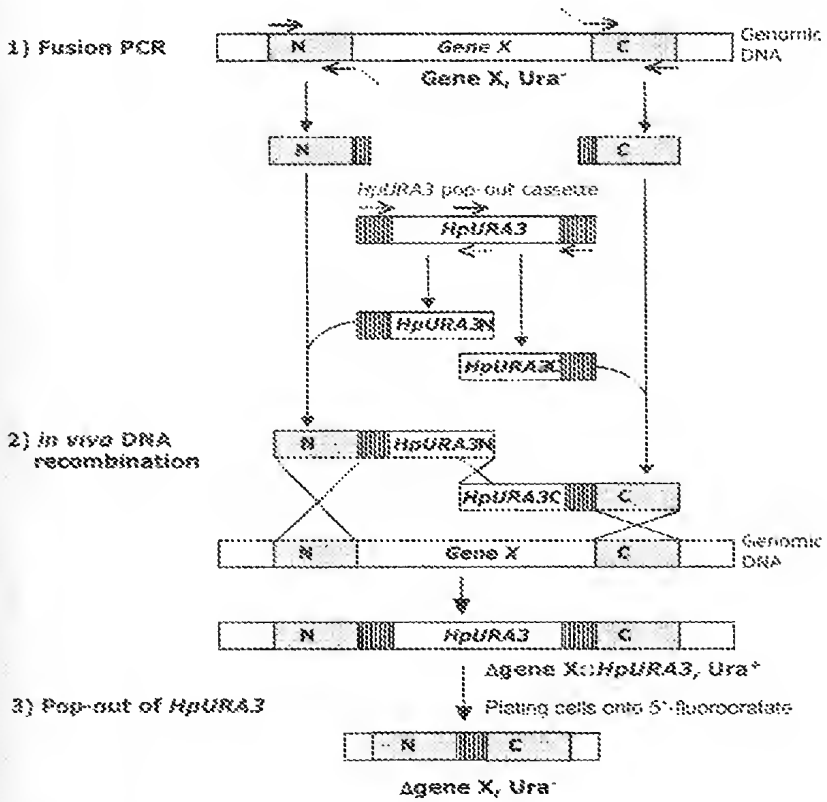


Fig. 6.6 Strategy of gene disruption using fusion PCR and in-vivo recombination in *H. polymorpha*. Step 1) Fusion-PCR. The N- and C-terminal fragments of "Gene X" are amplified and fused with the N- and C-fragments of the *HpURA3* pop-out cassette containing the overlapped internal sequence of *HpURA3* (100 bp) by PCR. Step 2) In-vivo DNA recombination. The two fusion-PCR products obtained are introduced into *H. polymorpha* cells and converted into one linear gene disruption cassette via in-vivo homologous recombination at the overlapped sequence. The double homologous crossover of the disruption cassette at the "Gene X" locus results in the disruption of "Gene X". Step 3) The *HpURA3* pop-out cassette can be removed to recover the auxotrophic marker for subsequent gene disruption or for subsequent transformation. A detailed procedure will be described elsewhere (MW Kim, HA Kang, unpublished results).

sporulation (Lahtchev 2002). Both CBS4732 (strain RB11 and its derivatives in particular) and DL-1 are employed in the production of recombinant products (Kang et al. 2001a; Gellissen 2000, 2002; Millet et al. 2002; Park et al. 2004; see also the forthcoming section of this chapter). In contrast to DL-1 strains, some sub-strains of CBS4732 are not easily applied to recombinatory methods, perhaps due to their high mitotic stability (Suckow and Gellissen 2002). For a selection of *H. polymorpha* strains and for protocols specific to parental strains, see Degelmann (2002) and the list in Appendix A6.1.

6.4.3
Plasmids and Available Elements

Expression and integration vectors in *H. polymorpha* are composed of prokaryotic and yeast DNA (Gellissen and Hollenberg 1997). Vectors are either supplied as circular plasmid or linearized and targeted to a specific genomic locus. Possible targets for homologous integration include the *MOX/TRP* locus (Agaphonov et al. 1995), an ARS sequence (Agaphonov et al. 1999; Sohn et al. 1996), the *URA3* gene (Brito et al. 1999), the *LEU2* gene (Agaphonov et al. 1999), the *GAP* promoter region (Heo et al. 2003), or the rDNA cluster (Klabunde et al. 2002, 2003). Clearly, the circular plasmids are not randomly integrated, but recombine with genomic sequences represented on the vector. This was recently shown with a particular vector harboring a *FMD* promoter/HBsAg fusion where recombination within the *FMD* gene was observed (U Dahlems, personal communication). It remains to be shown whether homologous recombination also takes place with vectors equipped with *MOX*, *TPS1* and other promoter elements.

Standard expression vectors and elements available for insertion into *H. polymorpha* as illustrated in Figures 6.7 and 6.8. (For detailed information on the various

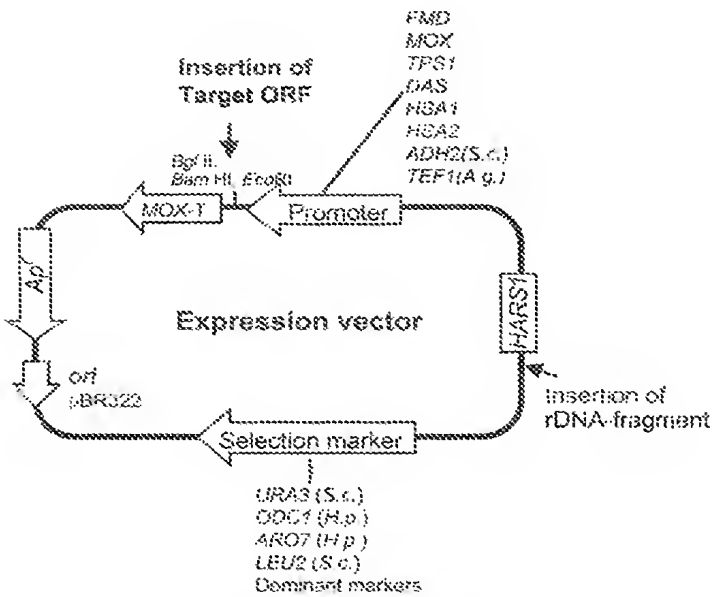
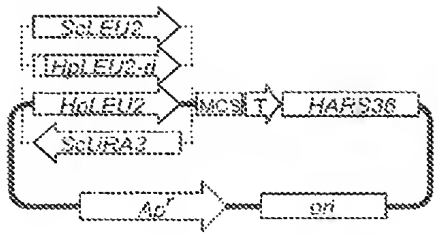
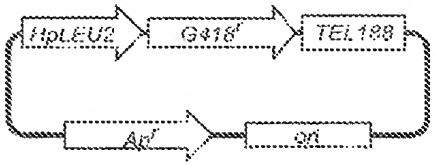


Fig. 6.7 General design of an expression vector for *H. polymorpha* RB11. A standard *H. polymorpha* expression vector contains the *ori* (origin of replication), a strong promoter, and a terminator (connected by a multiple cloning site for insertion of the foreign gene to be expressed), a selection marker from *H. polymorpha* or another yeast, and/or a selection marker for antibiotic resistance, a replication sequence (HARS), or alternatively a sequence for targeted integration into the genome sequence. For a list of elements available for insertion into a plasmid, see Table A6.2.

A. AMip series



B. pGLG series



C. pHACT-Hyl. series

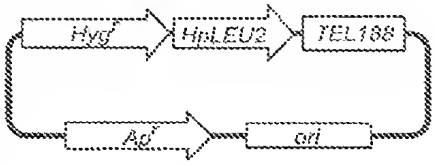


Fig. 6.8 Vectors designed for copy number-controlled gene integration in *H. polymorpha* DL-1 using auxotrophic selection markers (A) or antibiotic resistance markers (B and C). Refer to Kang et al. (2002) for more detailed information on these multiple integration vectors.

H. polymorpha expression platforms, see Suckow and Gellissen 2002; Kang et al. 2002; and Guengerich et al. 2004.) Plasmids that have been successfully developed for industrial use of RB11-based strains include pFPMT121 (for production of phytase) and a derivative of strain pMPT121 (for production of the anti-coagulant hirudin) (Suckow and Gellissen 2002) (Figure 6.7).

Plasmids AMpL1, AMpLD1, and AMpSU1 have been used as multiple integration systems based on the complementation of auxotrophic mutations, to elicit desired plasmid copy numbers in DL-1-derived recombinant strains (Figure 6.8A). When an appropriate mutant strain is transformed with one of these plasmids under selective conditions, transformants with plasmid integrated in low (1–2), moderate (6–9), or high (up to 100) copy numbers can be rapidly selected (Agaphonov et al. 1999). Another rapid and copy number-controlled selection system has been developed using antibiotic resistance markers (Figure 6.8B and C). The G418 and hygromycin B resistance cassettes were used as dominant selection markers, which allow selection of transformants on plates containing different concentrations of G418 (Sohn et al. 1999b) or hygromycin B (Kang et al. 2001b). Due to the strong correlation between antibiotic resistance and integration copy number, the selection of transformants with different copy numbers ranging from 1 to 50 can be easily

achieved. For a selection of *H. polymorpha* expression/integration vectors, refer to Degelmann (2002) and to the list provided in Appendix A6.1.

Plasmids harboring *CEN/ARS* (autonomously replicating sequences) elements typically support autonomous replication of plasmids within host cells. In *H. polymorpha*, centromeres have not yet been isolated. The established plasmids contain *HARS* elements, but not a *CEN* region. These plasmids may integrate into the host DNA over a number of generations, resulting in mitotically stable strains with as many as 100 plasmids in tandem repeats (Gellissen 2000; Kang et al. 2002). Notably, a high number of integrated copies is not always a pre-requisite for high-level protein production, especially in the case of secretory production. In one case, four copies of a *HARS* vector were sufficient to obtain efficient expression of *Schwanniomyces occidentalis* glucoamylase in *H. polymorpha* (Gellissen et al. 1992b). Other good examples are the secretory expression of human urinary-type plasminogen activator (u-PA) and human serum albumin (HSA) in *H. polymorpha*. In these cases, a single- or two-copy integration of the expression vector resulted in the maximum levels of recombinant u-PA or HSA secreted into culture supernatants (Kang et al. 2002).

Signal sequences may be fused to the target ORF for direct release of synthesized proteins into the media, or into a preselected cell compartment, such as the peroxisome, the vacuole, the ER, the mitochondrion, or the cell surface. Available signal sequences include the peroxisomal targeting signals PTS1 and PTS2 (van Dijk et al. 2000), the repressible acid phosphatase (*PHO1*) secretion leader sequence (Phongdara et al. 1998), a *S. occidentalis*-derived *GAM1* (van Dijk et al. 2000; Weydemann et al. 1995), and the *S. cerevisiae*-derived *MFα1* (Gellissen 2000). Glycosylphosphatidylinositol (GPI)-anchoring motifs derived from the GPI-anchored cell surface proteins, such as *HpSED1*, *HpGAS1*, *HpTIP1*, and *HpCWP1*, have been recently exploited to develop a cell surface display system in *H. polymorpha*. When the recombinant glucose oxidase (GOD) was expressed as a fusion protein to these anchoring motifs, most enzyme activity was detected at the cell surface (Kim et al. 2002).

One of the main advantages of heterologous gene expression in *H. polymorpha* is that this yeast has unusually strong promoters, the most widely employed of which are derived from genes of the methanol utilization pathway (see Figure 6.3). These promoters include elements derived from the methanol oxidase (*MOX*), formate dehydrogenase (*FMD*), and dihydroxyacetone synthase (*DHAS*) gene (Gellissen et al. 2000; van Dijk et al. 2000). Other available (but less frequently applied) regulative promoters are derived from inducible genes encoding enzymes involved in nitrate assimilation (e.g., *YNT1*, *YNI1*, *YNR1*, which can be induced by nitrate and repressed by ammonium) (Avila et al. 1998), or the enzyme acid phosphatase (the *PHO* promoter) (Phongdara et al. 1998, Baerends et al. 2002). Examples of constitutive promoters are *ACT* (Rang et al. 2001b), *GAP* (Heo et al. 2003), *PMA1* (Cox et al. 2000), and *TPS1* (Amuel et al. 2000). The *PMA1* promoter even competes with the outstanding *MOX* promoter in terms of high expression levels; *PMA1* is of interest in the co-expression of genes on industrial scale (Cox et al. 2000). The performance of the *TPS1* promoter is not linked to the use of a particular carbon source. In contrast to the constitutive promoters listed above, it can be applied at elevated temperatures, where its activity may be boosted even further (Amuel et al. 2000). The *H. polymorpha FLD1* gene en-

coding formaldehyde dehydrogenase has been characterized recently (Baerends et al. 2002). FLD1p is essential for the catabolism of methanol, and shows 82% sequence identity with the Fld1p protein from *P. pastoris* and 76% identity with Fld1p from *C. boidinii*. The *FLD1* promoter promises to be advantageous in that expression can be controlled at two levels: it is strongly induced under methylotrophic growth conditions, but shows moderate activity using primary amines as a nitrogen source. With these promising characteristics, the *FLD1* promoter is expected to augment the existing *H. polymorpha* promoters (Baerends et al. 2002). The *GAP* promoter also showed a higher specific production rate and required a much simpler fermentation process than the *MOX* promoter-based HSA production system, implying that the *GAP* promoter can be a practical alternative of the *MOX* promoter in the large-scale production of some recombinant proteins (Heo et al. 2003).

6.5
Product and Process Examples

We now provide a short summary of *H. polymorpha*-based processes. Here, a few industrially relevant examples are only briefly summarized; for a more detailed description of fermentation and purification procedures, the reader is referred to Jenzelewski (2002) and to Chapter 14 of this book.

Once stable recombinant integrants have been generated, production strain candidates are identified from a background of nonproducers or strains of low or impaired productivity. The subsequent design of a fermentation procedure greatly depends on the characteristics of the host cell, the intended routing of the recombinant gene product and, most importantly, on the promoter elements used. The commonly used culture media are based on simple synthetic components. They contain trace metal ions and adequate nitrogen sources, which are required for efficient gene expression and cell yield, but no proteins. The total fermentation time varies between 60 and 150 hours. Due to the inherent versatile characteristics of the two methanol-inducible promoters, fermentation modes vary for the most part in the supplemented carbon source: glycerol, methanol, glucose, and combinations thereof may be selected. The ability to achieve high yields of a recombinant product expressed from a methanol pathway promoter without the addition of methanol is a unique feature of the *H. polymorpha* system (see Section 6.1; Gellissen 2000; Suckow and Gellissen 2002). In contrast, activation of these promoters in the related yeast *P. pastoris* is strictly dependent on the presence of methanol (Cregg 1999).

In processes for secretory heterologous proteins, a "one-carbon source" mode is usually employed, supplementing the culture medium only with glycerol. A hirudin production process may serve as an example for this fermentation mode. In this process, a strain was employed that harbors 40 copies of an expression cassette for an *MFα1* pre-pro-sequence/hirudin fusion gene under control of the *MOX* promoter (Weydemann et al. 1995; Avgerinos et al. 2001; Bartelsen et al. 2002). Hirudin production was promoted by reducing the initial glycerol concentration and maintaining it on a suitable level by a pO_2-controlled addition of the carbon source. The fer-

mentation is started with 3% (w/v) glycerol. After consumption of the carbon source (after 25 h), the pO_2-controlled feeding mode is initiated and this results in a glycerol concentration of between 0.05 and 0.3% (w/v) (derepression of the *MOX* promoter). The fermentation run is terminated after 36 h of derepression (total fermentation time 72 h). The broth is then harvested and the secreted product purified from the supernatant by a sequence of ultrafiltration, ion exchange, and gel filtration steps.

In the case of HBsAg production, a "two carbon source" fermentation mode was used (see Chapter 15). The producer strain harbors high copy numbers of an expression cassette, with the coding sequence for the small surface antigen (S-antigen) under control of methanol pathway promoters. The selected strain is fermented on a 50-L scale. The product-containing cells are generated via a two-fermentor cascade, consisting of a 5-L seed inoculating the 50-L main fermentor. The initial steps of fermentation closely follow those described for the production of hirudin. Initially, cultivation is performed with a glycerol feeding in a fed-batch mode, to be followed by subsequent semi-continuous glycerol feeding controlled by the dissolved oxygen level in the culture broth. This derepression phase is then followed by batchwise feeding with methanol in the final fermentation mode. The product concentration increases to amounts in the multigram range. The product consists of a lipoprotein particle in which the recombinant HBsAg is inserted into host-derived membranes. As noted in Section 6.1, the addition of methanol also serves for the proliferation of organelles and consequently for the synthesis and proliferation of membranes. Methanol is thus needed in this case to provide a high yield and balanced co-production of both components of the particle (Schaefer et al. 2001, 2002; see also Chapter 15). For downstream processing, the harvested cells are disrupted and the particles purified in a multi-step procedure that includes adsorption of a debris-free extract to a matrix and the subsequent application of a sequence of ion exchange, ultra-filtration, gel filtration, and ultra-centrifugation steps as detailed by Schaefer et al. (2001, 2002), and in Chapter 15 for recombinant hepatitis B vaccines.

For the production of phytase, *H. polymorpha* has been used in a particularly efficient process (Mayer et al. 1999; Papendieck et al. 2002), a pre-requisite for an economically competitive production of a technical enzyme. In this development all steps and components of strain generation, fermentation, and purification are dictated by a rationale of efficiency and cost-effectiveness. This also applies to the definition of fermentation process using glucose as the main carbon source.

A strain was generated in which the phytase sequence is under control of the *FMD* promoter. Subsequent supertransformation yielded strains with up to 120 copies of the heterologous DNA, thus enabling a gene dosage-dependent high productivity. A fermentation procedure was then developed to achieve high levels of enzyme production. Significantly, it was found that the use of glycerol as the main carbon source was not required in this case, but could be substituted by low-cost glucose. The active status of the *FMD*-promoter was maintained by glucose starvation (fermentation with minimal levels of continuously fed glucose). At a 2000-L scale, fermentation with glucose as the sole carbon source led to high product yields and an 80% reduction in raw material costs compared to glycerol-based fermentations (Mayer et al. 1999; Papendieck et al. 2002). Strains were found to produce the recom-

Tab. 6.5 *H. polymorpha*-based products (selection).

Category	Product	Status	Brand name	Reference
Pharmaceuticals	HBsAg (*adr*)	Launched	HepaVax Gene	Schaefer et al. (2002)
	HBsAg (*adw*)	Launched	AgB	Schaefer et al. (2002)
	Insulin	Launched	Wosulin	
	IFNα-2a	Process transfer		Müller H et al. (2001)
	HSA	Pilot scale completed		Heo et al. (2003)
	EGF	Lab scale completed		Heo et al. (2002)
Food additive	hexose oxidase	Launched	Grindamyl Surebake	Cook and Thygesen (2003)
Food additive	phytase	Registration		Mayer et al. (1999)
Enzymes	Levansucrase	Lab scale completed		Park et al. (2004)

binant phytase at levels ranging up to 13.5 g L^{-1} (Mayer et al. 1999). The secreted product is purified through a series of steps, including flocculation centrifugation, dead-end filtration, and a final ultra-filtration yielding a high-quality, highly concentrated product at a recovery rate of up to 92%.

A short outline of product recovery and downstream processing was provided for the previous examples of two secreted and one intracellular product. An individual procedure must be defined for each process developed. Especially in the case of secreted compounds, the fermentation and primary product recovery are intimately linked, and this interface of upstream and downstream processing is often the objective of successful integrated bioprocess development (Curvers et al. 2001; Thömmes et al. 2001). A typical example of this is the production of aprotinin variants in *H. polymorpha* (Zurek et al. 1996).

A selection of *H. polymorpha*-derived products is listed in Table 6.5.

6.6 Future Directions and Conclusion

6.6.1 Limitations of the *H. polymorpha*-based Expression Platform

Despite the most favorable characteristics of the *H. polymorpha*-based platform for application in heterologous gene expression, problems and limitations may be encountered in particular strain and product developments, as is similarly (and more frequently) seen in other yeast systems. These limitations include overglycosylation (Agaphonov et al. 2001), retention within the ER (Agaphonov et al. 2002), poor secre-

Tab. 6.6 Improvement of the expression performance of *H. polymorpha* production strains by the co-expression or deletion of other genes.

Gene expressed	Problem encountered	Co-expressed (+) or deleted (–) gene
IFNα-2a	Incorrect pre-pro-cleavage	*KEX2* (+)
Enzyme, IFNγ	Overglycosylation impaired secretion	*CMK2*; *CNE1* (+)
EGF	C-terminal truncation	*KEX1* (–)

tion, impaired processing (Müller H et al. 2001; Gellissen et al. 2002), and proteolytic degradation (Suckow and Gellissen 2002). A possible strategy to overcome these limitations is to identify genes and gene products that may, upon disruption or co-expression, positively influence the performance of respective strains. This has been applied successfully in several cases. For example, the deletion of the *KEX1* gene, coding for carboxypeptidase κ, led to a significant improvement in the quality of recombinant human epidermal growth factor (hEGF) secreted from *H. polymorpha* by decreasing the generation of C-terminally truncated hEGF form (Heo et al. 2003). Among others, co-expression of the *S. cerevisiae*-derived *KEX2* gene provided a greatly improved processing of a IFNα-2a pre-pro-sequence in *H. polymorpha* in which production of predominantly N-terminally extended molecules had been observed previously (Müller H et al. 2001; Gellissen et al. 2002). In other examples, the co-expression of the *S. cerevisiae*-derived *CMK2* or the *H. polymorpha CNE1* (calnexin) gene led to an improved secretion and a reduction in overglycosylation of a secreted enzyme and a cytokine (Table 6.6).

6.6.2
Impact of Functional Genomics on Development of the *H. polymorpha* RBTI-based Expression Platform

Several approaches have been initiated to identify *H. polymorpha* genes that may be manipulated to effect a positive influence on the performance of particular production strains. Examples include identification of the *PMR1* gene (Kang et al. 1998) and glycosylation genes (Gnengerich et al. 2004; Kim et al. 2002). With the completion of genome sequencing, transcriptome, proteome analysis and other related technologies are all now feasible and enable more systematic approaches to be introduced.

In a first approach, genes will be identified that are involved in methanol metabolism, peroxisome homeostasis, protein glycosylation, secretion, and cell wall integrity. These tasks are executed within a cooperative effort with partners in Russia, Ukraine, The Netherlands and Germany, and funded by INTAS (INTAS 2001–0583). For the identification of these genes, linear DNA fragments harboring reporter genes are used for random integration, thereby generating mutants. By using this random integration (RALF) approach (van Dijk et al. 2001), certain genes of potential impact for relevant gene expression functions may be disrupted and identified by sequencing the region adjoining the integration site and comparing the deduced sequence with the genome data. By applying a selection of suitable reporter proteins and a

range of certain growth conditions, the generated strains can be screened for genes which might have an impact on the functions mentioned above.

For transcriptome analysis, *H. polymorpha* cDNA microarrays are being generated in cooperation with KRIBB (Korea Research Institute of Bioscience and Biotechnology, Daejeon, Korea), funded by the Korean Ministry of Science and Technology (Microbial Genomics and Applications R& D Program). As an initial trial, a partial genome cDNA chip spotted with 382 ORFs of *H. polymorpha* was constructed. Each ORF was PCR-amplified using gene-specific primer sets, of which the forward primers have a 5'-aminolink. The PCR products were printed in duplicate onto the aldehyde-coated slide glasses to link only the coding strands to the surface of the slide via covalent coupling between amine and aldehyde groups. The partial DNA chip was used to analyze differential expression profiling of *H. polymorpha* cells cultivated under defined environmental conditions (Oh et al. 2004). At present, the whole-genome cDNA microarrays of *H. polymorpha* are available, and these have been constructed using the same strategy as applied to fabrication of the partial cDNA chip. It is expected that these whole-genome DNA microarrays will provide a powerful tool for a genome-wide transcriptional profiling of *H. polymorpha* under defined genetic and physiological conditions. This will generate invaluable information for pathway engineering and process optimization in exploiting *H. polymorpha* as a cell factory.

A third project which is to be started soon is a comprehensive analysis of the proteome of recombinant *H. polymorpha* production strains in correlation to specific products, secretion efficiency, and other characteristics. The extraction of defined proteins from two-dimensional SDS gels and mass spectrometric analysis of proteolytic fragments will lead to the identification of proteins and their respective genes, with a potential impact on the performance of the *H. polymorpha* expression platform. The availability of the complete genome and various 'omic's approaches surely facilitates extensive exploration of interesting genes and strong promoters, and will thus further the development of expression systems to supplement the strong platform that already exists for *H. polymorpha*.

Yeasts have only come under intense molecular study during the past few decades, and yeast-derived recombinant products have been already developed, ranging from technical enzymes and anticoagulants (hirudin and saratin) (Avgerinos et al. 2001; Sohn et al. 2001; Barnes et al. 2001; Bartelsen et al. 2002) to vaccines such as hepatitis B (Schaefer et al. 2001; Schaefer et al. 2002). *S. cerevisiae* is well characterized, being the first organism in which recombinant vaccines were developed, and the first eukaryotic organism to have its genome entirely sequenced. However, *H. polymorpha* has several advantages over *S. cerevisiae*, including strong and tightly regulated promoters, the lack of hyperallergenic structures on target proteins, capabilities of dense growth on simple media, stability of expression plasmids, and high frequency of nonhomologous recombination. Moreover, *H. polymorpha* has the endogenous capacity for prolyl 4-hydroxylation, which is essential for the function and folding of certain recombinant proteins, such as gelatin (de Bruin et al. 2002). This post-translational modification is generally known to be absent in microbial eukaryotic systems. Consequently, *H. polymorpha* holds a secure place in representing the methylotrophic yeasts as an alternative system for heterologous gene expression.

Appendix

Tab. A6.1 Selection of *H. polymorpha* host strains.

Strain	Genotype	Phenotype	Source
Parental strain			
DL-1	wild-type (NRRL-Y-7560, ATCC26012)		Levine and Cooney (1973)
Auxotrophic strains			
DL-1-L	*leu2*	Leu$^-$	Sohn et al. (1996)
uDL10	*leu2 ura3*	Leu$^-$Ura$^-$	KRIBB
DL-1dU	*leu2 Δura3::lacZ*	Leu$^-$Ura$^-$	KRIBB
DL1A-A	*leu2 Δade2*	Leu$^-$Ade$^-$	CRC
DL1A-L	*Δade2 Δleu2::ADE2*	Leu$^+$	CRC
DL1A-U	*leu2 Δade2 Δura3::ADE2*	Leu$^-$Ura$^-$	CRC
Protease-deficient strains			
uDLB11	*leu2 ura3 Δpep4::lacZ*	Leu$^-$ Ura$^-$ Pep4$^-$	KRIBB
uDLB12	*leu2 ura3 Δprc1::lacZ*	Leu$^-$Ura$^-$ Prc1$^-$	KRIBB
uDLB13	*leu2 ura3 Δkex1::lacZ*	Leu$^-$ Ura$^-$ Kex1$^-$	KRIBB
uDLB14	*leu2 ura3 Δpep4::lacZ Δprc1::lacZ*	Leu$^-$Ura$^-$Pep4$^-$ Prc1$^-$	KRIBB
uDLB15	*leu2 ura3 Δpep4::lacZ Δkex1::lacZ*	Leu$^-$Ura$^-$Pep4$^-$ Kex1$^-$	KRIBB
uDLB16	*leu2 ura3 Δprc1::lacZ Δkex1::lacZ*	Leu$^-$Ura$^-$ Prc$^-$ Kex1$^-$	KRIBB
uDLB17	*leu2 ura3 Δpep4::lacZ Δprc1::lacZ Δkex1::lacZ*	Leu$^-$Ura$^-$Pep4$^-$Prc1$^-$ Kex1$^-$	KRIBB
Methanol utilization-deficient strains			
DU12	*leu2 Δmox-trp3::ScLEU2*	Mox$^-$Trp$^+$	CRC
DL1-LAM	*leu2 Δmox*	Leu$^+$Mox$^-$	CRC
Parental strain			
NCYC495	wild-type (CBS1976, ATAA14754, NRRL-Y-1798,VKM-Y-1397)		Wickerham (1951)
L1	*leu1-1**	Leu$^-$	Gleeson et al. (1986)
A11	*ade11-1*	Ade$^-$	Parpinello et al. (1998)
M6	*met6-1*	Met$^-$	Parpinello et al. (1998)
NCYC495	*leu1-1**	Leu$^-$	Brito et al. (1999)
Nitrate assimilation-related strains			
NAG1995	*Δynr1::URA3, leu1-1**	Ynr1$^-$ Leu$^-$	Avila et al. (1995)
NAG1996	*Δyni1::URA3, leu1-1**	Yni1$^-$ Leu$^-$	Brito et al. (1996)
NAG997	*Δyni1::URA3, leu1-1**	Yni1$^-$ Leu$^-$	Pérez et al. (1997)
NAG998	*Δyna1::URA3, leu1-1**	Yna1$^-$ Leu$^-$	Avila et al. (1998)
NAG2001	*yna2::URA3, leu1-1**	Yna2$^-$	Avila and Siverio (unpublished)
Parental strain			
CBS4732	wild-type (CCY38-22-2, ATCC34438, NRRL-Y-5445)		Morais and Maia (1959)
L89	*ura3-1 (ade1)*	Ura$^-$	Roggenkamp et al. (1986)
RB11	*ura3-1*	Ura$^-$	Weydemann et al. (1995)
RB12	*ura3 leu1-1**	Ura$^-$Leu$^-$	Rhein Biotech, unpublished
RB13	*ura3 leu1-1* mox*	Ura$^-$Leu$^-$Mox$^-$	Rhein Biotech, unpublished

Tab. A6.1 (continued)

Strain	Genotype	Phenotype	Source
RB14	*ura3 mox*	Ura$^-$ Mox$^-$	Rhein Biotech, unpublished
RB15	*leu1–1* mox*	Leu$^-$ Mox$^-$	Rhein Biotech, unpublished
RB17	*haro7*	Tyr$^-$	Kruppmann et al. (2000)
RC296	*ade*	Ade$^-$	Rhein Biotech, unpublished
A16	*leu2*	Leu$^-$	
	trp3 mox	Trp$^-$ Mox$^-$	Veale et al. (1992)
1B	*ade2–88 leu2–2*	Ade$^-$ Leu$^-$	Bogdanova et al. (1998)
1-HP065	*ade2–88 ura2–1 met 4–220*	Ade$^-$ Leu$^-$ Met$^-$	Marmazzu et al. (1997)
14C	*leu2–2 cat1–14*	Leu$^-$ Cat$^-$	Lahtchev (2002)
5C-HP156	*ade2–88*	Ade$^-$	Lahtchev (2002)
8V	*leu2*	Leu$^-$	Agaphanov et al. (1995)

* *leu1–1* and *leu2* correspond to the same gene.

Tab. A6.2 Selection of *H. polymorpha* expression/integration vectors.

Plasmid	Expression cassette	Replication sequence	Selection marker	Integrated copy number	References/comments
DL-1 based plasmids					
AMIpL1	No promoter; Terminator from an unknown gene	*HARS36*	*HpLEU2*	1–2	Agaphonov et al. (1999) Multiple cloning sites for insertion of expression cassettes
AMIpLD1	No promoter; Terminator from an unknown gene	*HARS36*	*HpLEU2-d*	1–100	Agaphonov et al. (1999) Multiple cloning sites for insertion of expression cassettes
AMIpSL1	No promoter; Terminator from an unknown gene	*HARS36*	*ScLEU2*	6–9	Agaphonov et al. (1999) Multiple cloning sites for insertion of expression cassettes
AMIpSU1	No promoter; Terminator from an unknown gene	*HARS36*	*ScURA3*	30–50	Agaphonov et al. (1999) Multiple cloning sites for insertion of expression cassettes
pGLC61	No promoter; No terminator	*TEL188*	*HpLEU2/G418*^{r}	1–50	Sohn et al. (1999) *Not*I and a *Bam*HI sites for insertion of expression cassettes
pHACT90 HyL	No promoter; No terminator	*TEL188*	*HpLEU2/Hyg*^{r}	1–25	Kang et al. (2001) A *Not*I site for insertion of expression cassettes

Tab. A6.2 (continued)

Plasmid	Expression cassette	Replication sequence	Selection marker	Integrated copy number	References/comments
pBR1-based plasmids					
pMPT121	*MOX*-promoter; *MOX*-terminator	*HARS1* (oppositely oriented than in pFPMT121 and pTPSPMT)	*ScURA3*	30–60	Gellissen and Hollenberg (1997) *EcoR*I, *BamH*I, *Bgl*I, sites for insertion of ORFs
pFPMT121	*FMD*-promoter; *MOX*-terminator	*HARS1*	*ScURA3*	30–60	Zurek et al. (1996) *EcoR*I, *BamH*I, *Bgl*I, sites for insertion of ORFs
pTPSIMT121	*TPS1*-promoter *MOX*-terminator	*HARS1*	*ScURA3*	30–60	Rhein Biotech, unpublished; *Eco*RI, *Bam*HI, *Bgl*II, sites for insertion of ORFs
pB14	*FMD*-promoter; *MOX*-terminator	*HARS1*	*ScURA3*	30–60	Rhein Biotech, unpublished; *Eco*RI, *Bam*HI *Bgl*II, sites for insertion of ORFs; no ampR gene
pB14-TPS1	*TPS1*-promoter; *MOX*-terminator	*HARS1*	*ScURA3*	30–60	Rhein Biotech, unpublished; *Eco*RI, *Bam*HI, *Bgl*II, sites for insertion of ORFs; no ampR
pB14-LEU2	*FMD*-promoter; *MOX*-terminator	*HARS1*	*ScLEU2*	30–60	Rhein Biotech, unpublished; *Eco*RI, *Bam*HI, *Bgl*II sites for insertion of ORFs; no ampR
pM1	No promoter *MOX*-terminator	*HARS1*	*ScURA3*	n.d	Amuel et al. (2000) Multiple cloning sites for insertion of expression cassettes
pSK92	*FMD*-promoter; *MOX*-terminator	*HARS1*	*HARO7*	1–5	Krappmann et al. (2000) Rhein Biotech, unpublished, *Eco*RI, *Bam*HI, *Bgl*II, sites for insertion of ORFs

Tab. A6.2 (continued)

Plasmid	Expression cassette	Replication sequence	Selection marker	Integrated copy number	References/comments
NCYC495-based plasmids					
pHIPA4	*AOX*-promoter *AMO*-terminator	No replication sequence	*HpPUR7*/Amp	n.d.	Haan et al. (2002)
pHIPX4	*AOX*-promoter *AMO*-terminator	No replication sequence*	*ScLEU2*/Kan	n.d.	Gietl et al. (1994)
pHIPX5	*AMO*-promoter *AMO*-terminator	No replication sequence*	*ScLEU2*/Kan	n.d.	Kiel et al. (1995)
pHIPX6	*PEX3*-promoter *AMO*-terminator	No replication sequence*, **	*ScLEU2*/Kan	n.d.	Kiel et al. (1995)
pHIPX7	*TEF1*-promoter *AMO*-terminator	No replication sequence*	*ScLEU2*/Kan	n.d.	Baerends et al. (1997)
pHIPX8	*TEF2*-promoter *AMO*-terminator	No replication sequence*	*ScLEU2*/Kan	n.d.	M. Veenhuis, unpublished
pHIPZ4	*AOX*-promoter *AMO*-terminator	No replication sequence	Zeocin/Amp	n.d.	Salomons et al. (2000)
pREMI-Z	*ScTEF1*-promoter/ EM7 synthetic promoter/ *Sc*CYC-terminator	No replication sequence	Zeocin	n.d.	Van Dijk et al. (2002)
pYT3	No promoter No terminator	*C-ARS* *	*ScLEU2*/Amp	n.d.	Tan et al. (1995)
pHS5	No promoter No terminator *LacZ*α	No replication sequence*	*ScLEU2*/Amp	n.d.	M. Veenhuis, unpublished
pHS6	No promoter No terminator *LacZ*α	*HARS1**	ScLEU2/Amp	n.d.	M. Veenhuis, unpublished
pBSK-LEU2-Ca	No promoter No terminator	No replication sequence	CaLEU2/Amp	n.d.	M. Veenhuis, unpublished
pH11	No promoter No terminator *LacZ*α	No replication sequence	*HpURA3*/Amp	n.d.	Kiel et al. (1999)

Key: Ca, Candida albicans; Hp, Hansenula polymorpha; Sc, Saccharomyces cerevisiae.

* These plasmids contain the *S. cerevisiae LEU2* gene, which harbours a cryptic HARS. As a consequence, pHIPX- and pHS-type plasmids replicate – albeit rather poorly – in *H. polymorpha* NCYC495. Addition of the C-ARS or HARS1 regions results in good replicating plasmids.

** pHIPX6 contains the *H. polymorpha PEX3* promoter. This fragment contains a HARS, and allows use as a replicating plasmid to express genes at rather low levels.

References

Agaphonov MO, Bibilov MY, Ter-Avanesyan MD, Smirnov VN (1995) Disruption-displacement approach for the targeted integration of foreign genes in *Hansenula polymorpha*. Yeast 11: 1241–1247

Agaphonov MO, Deev AV, Kim S-Y, Sohn J-H, Choi E-S, Ter-Avanesyan MD (2003) A novel approach to isolation and functional characterization of genomic DNA sequences from the methylotrophic yeast *Hansenula polymorpha*. Mol Biol 37: 74–80

Agaphonov MO, Packeiser AN, Cherkebnova MB, Choi E-S, Ter-Avanesyan MD (2001) Mutation of the homologue of GDP-mannose pyrophosphorylase alters cell wall structure, protein glycosylation and secretion in *Hansenula polymorpha*. Yeast 18: 391–402

Agaphonov MO, Romanos NV, Trushkina PM, Smirnov VN, Ter-Avanesyan MD (2002) Aggregation and retention of human urokinase-type plasminogen activator in the yeast endoplasmic reticulum. BMC Mol Biol 3: 15–22

Agaphonov M, Trushkina PM, Sohn JS, Choi ES, Rible SK, Ter-Avanesyan MD (1999) Vectors for rapid selection of integrants with different plasmid copy numbers in the yeast *Hansenula polymorpha* DL-1. Yeast 15: 541–551

Amuel C, Gellissen G, Hollenberg CP, Suckow M (2000) Analysis of heat shock promoters in *Hansenula polymorpha*: TPS1, a novel element for heterologous gene expression. Biotechnol Bioprocess Eng 5: 247–252

Avgerinos GC, Turner BG, Gorelick MD, Papendieck A, Weydemann U, Gellissen G (2001) Production and clinical development of a *Hansenula polymorpha*-derived PEGylated hirudin. Semin Thromb Hemost 27: 357–371

Avila J, González C, Brito N, Siverio JM (1998) Clustering of the *YNA1* gene encoding a $Zn(II)_2Cys_6$ transcriptional factor in the yeast *Hansenula polymorpha* with the nitrate assimilation genes *YNT1*, *YNI1* and *YNR1*, and its involvement in their transcriptional activation. Biochem J 335: 647–652

Avila J, Pérez MD, Brito N, González C, Siverio JM (1995) Cloning and disruption of the *YNR1* gene encoding the nitrate reductase apoenzyme of the yeast *Hansenula polymorpha*. FEBS Lett 366: 137–142

Baerends RJS, Salomons FA, Kiel JAKW, van der Klei IJ, Veenhuis M (1997) Deviant Pex3p levels affect normal peroxisome formation in *Hansenula polymorpha*: a sharp increase of the protein level induces the proliferation of numerous, small protein import-competent peroxisomes. Yeast 13: 1449–1463

Baerends RJS, Sulter GJ, Jeffries TW, Cregg JM, Veenhuis M (2002) Molecular characterization of the *Hansenula polymorpha PLD1* gene encoding formaldehyde dehydrogenase. Yeast 19: 37–42

Barnes CS, Kraffe B, Frech M, Hofmann UR, Papendieck A, Dahlems U, Gellissen G, Hovlaeris MP (2001) Production and characterization of saratin, an inhibitor of von Willebrand factor-dependent platelet adhesion to collagen. Semin Thromb Hemost 27: 337–347

Barnett JA, Payne RW, Yarrow D (2000) Yeasts: Characteristics and Identification, 3rd edition. Cambridge University Press, Cambridge, UK

Barr PJ, Brake AJ, Valenzuela P (Eds) Yeast genetic engineering. Butterworth, Boston, MA

Barteldsen O, Barnes CS, Gellissen G (2002) Production of anticoagulants in *Hansenula polymorpha*. In: Gellissen G (Ed) *Hansenula polymorpha* – biology and applications. Wiley-VCH, Weinheim, pp 211–228

Birney E, Durbin R (2000) Using GeneWise in the *Drosophila* annotation experiment. Genome Res 10: 547–548

Blandin G, Llorente B, Malpertuy A, Wincker P, Artiguenave F, Dujon B (2000) Genomic exploration of the hemiascomycetous yeasts: 13. *Pichia angusta*. FEBS Lett 487: 76–81

Bogdanova AI, Kustikova OS, Agaphonov MO, Ter-Avanesyan MD (1998) Sequences of *Saccharomyces cerevisiae* 2μm DNA improving plasmid partitioning in *Hansenula polymorpha*. Yeast 14: 1–9

Brito N, Avila J, Pérez MD, González C, Siverio JM (1996) The genes *YNI1* and *YNR1* encoding nitrite reductase and nitrate reductase respectively in the yeast *Hansenula polymorpha*, are clustered and co-ordinately regulated. Biochem J 317: 89–95

Brito N, Pérez MD, Perdomo G, González C, García-Loco P, Siverio JM (1999) A set of *Hansenula polymorpha* integrative vectors to construct *lacZ* fusions. Appl Microbiol Biotechnol 53: 23–29

Bretthauer RK, Castellano FJ (1999) Glycosylation of *Pichia pastoris*-derived proteins. Biotechnol Appl Biochem 30: 193–200

Callewaert N, Laroy W, Cadirgi H, Geysens S, Saelens X, Jou WM, Contreras R (2001) Use of HDEL-tagged *Trichoderma reesei* mannosyl oligosaccharide 1,2-α-D-mannosidase for N-glycan engineering in *Pichia pastoris*. FEBS Lett 503: 173–178

Chiba Y, Suzuki M, Yoshida S, Yoshida A, Ikenaga H, Takeuchi M, Jigami Y, Ichishima E (1998) Production of human compatible high mannose-type ($Man_5GlcNAc_2$) sugar chains in *Saccharomyces cerevisiae*. J Biol Chem 273: 26298–26304

Cook MW, Tryggseth HV (2003) Safety evaluation of a hexose oxidase expressed in *Hansenula polymorpha*. Food Chem Toxicol 41: 523–529

Cox H, Mead D, Sudbery, Eland M, Evans L (2000) Constitutive expression of recombinant proteins in the methylotrophic yeast *Hansenula polymorpha* using the *PMA1* promoter. Yeast 16: 1191–1203

Cregg JM (1999) Expression in the methylotrophic yeast *Pichia pastoris*. In: Fernandez JM, Hoeffler JP (Eds) Gene expression systems: using nature for the art of expression. Academic Press, San Diego, pp 157–191

Curvers S, Brinkus P, Klauser T, Weuster-Botz D, Takors R, Wandrey C (2001) Human chymotrypsinogen B production with *Pichia pastoris* by integrated development of fermentation and downstream processing Part I. Fermentation. Biotechnol Progr 17: 495–502

De Bruin EC, Wenten MW, Laane C, de Wolf FA (2002) Endogenous propyl 4-hydroxylation in *Hansenula polymorpha* and its use for the production of hydroxylated recombinant gelatin. FEMS Yeast 1: 291–298

Degelmann A (2002) Methods. In: Gellissen G (Ed) *Hansenula polymorpha* – biology and applications. Wiley-VCH, Weinheim, pp 285–335

Degelmann A, Müller F, Sieber H, Jenzelewski V, Suckow M, Strasser AWM, Gellissen G (2002) Strain and process development for the production of human cytokines in *Hansenula polymorpha*. FEMS Yeast Res 2: 349–361

Dubois T, Roggenkamp R (1992) Targeting signal of the peroxisomal catalase in the methylotrophic yeast *Hansenula polymorpha*. FEBS Lett 303: 113–116

Feldmann, H (Ed) (2000) Génolevures. Genomic exploration of the hemiascomycetous yeasts. FEBS Lett 487: 1–150

Gellissen G (2000) Heterologous protein production in methylotrophic yeasts. Appl Microbiol Biotechnol 54: 741–750

Gellissen G (2002) *Hansenula polymorpha* – biology and applications. Wiley-VCH, Weinheim

Gellissen G, Hollenberg CP (1997) Application of yeasts in gene expression studies: a comparison of *Saccharomyces cerevisiae*, *Hansenula polymorpha* and *Kluyveromyces lactis* – a review. Gene 190: 87–97

Gellissen G, Veenhuis M (2001) The methylotrophic yeast *Hansenula polymorpha*: its use in fundamental research and as a cell factory. Yeast 18: v–iii

Gellissen G, Janowicz ZA, Weydemann U, Melber K, Strasser AWM, Hollenberg CP (1992a) High-level expression of foreign genes in *Hansenula polymorpha*. Biotechnol Adv 10: 179–189

Gellissen G, Melber K, Janowicz SA, Dahlems U, Weydemann U, Piontek M, Strasser AWM, Hollenberg CP (1992b) Heterologous protein expression in yeast. Antonie van Leeuwenhoek 62: 79–93

Gellissen G, Müller F, Sieber H, Dierk A, Jenzelewski V, Degelmann A, Strasser AWM (2002) Production of cytokines in *Hansenula polymorpha*. In: Gellissen G (Ed) *Hansenula polymorpha* – biology and applications. Wiley-VCH, Weinheim, pp 229–254

Gellissen G, Piontek M, Dahlems U, Jenzelewski V, Gavagan JE, DiCosimo R, Anton DL, Janowicz SA (1996) Recombinant *Hansenula polymorpha* as a biocatalyst: coexpression of the spinach glycolate oxidase (*GO*) and the *S. cerevisiae* catalase T (*CTT1*) gene. Appl Microbiol Biotechnol 46: 46–54

Gemmill TR, Trimble RB (1999) Overview of N-linked and O-linked oligosaccharide structures found in various yeast species. Biochim Biophys Acta 1426: 227–237

Gietl C, Faber KN, van der Klei IJ, Veenhuis M (1994) Mutational analysis of the N-terminal topogenic signal of watermelon glyoxysomal malate dehydrogenase using the heterologous host *Hansenula polymorpha*. Proc Natl Acad Sci USA 91: 3151–3155

Gleeson MA, Ortori GS, Sudbery PE (1986) Transformation of the methylotrophic yeast *Hansenula polymorpha*. J Gen Microbiol 132: 3459–3465

Goffeau A (1996) 1996: A vintage year for Yeast and yeast. Yeast 2: 1603–1606

Gonzalez C, Perdomo G, Tejera P, Brito N, Siverio JM (1999) One-step, PCR-mediated, gene disruption in the yeast *Hansenula polymorpha*. Yeast 15: 1323–1329

Gutgerich L, Kang HA, Bebie B, Gellissen G, Sieckow M (2004) A platform for heterologous gene expression based on the methylotrophic yeast *Hansenula polymorpha*. In: Flack U (Ed) The mycota II. Genetics and Biotechnology. 2nd edition. Springer-Verlag, Heidelberg, pp 273–287

Haan GJ, van Dijk R, Kiel JAKW Veenhuis M (2002) Characterization of the *Hansenula polymorpha POT1* gene and its use as selectable marker for targeted chromosomal integration. FEMS Yeast Res 2: 17–24

Hamilton SR, Bobrowicz P, Bobrowicz B, Davidson RC, Li H, Mitchell T, Nett JH, Rausch S, Stadheim TA, Wischnewsk H, Wildt S, Gerngross TU (2003) Production of complex human glycoproteins in yeast. Science 301: 1244–1246

Hazeu W, de Bruyn JC, Bos P (1972) Methanol assimilation by yeasts. Arch Mikrobiol 87: 185–188

Heo JH, Won HS, Kang HA, Rhee SK, Chung BH (2002) Purification of recombinant human epidermal growth factor secreted from the methylotrophic yeast *Hansenula polymorpha*. Protein Expr Purif 24: 117–122

Heo JH, Hong WK, Choi EY, Kim MW, Kim JY, Kim CH, Rhee SE, Kang HA (2003) Properties of the *Hansenula polymorpha*-derived constitutive *GAP* promoter, assessed using an *HSA* reporter gene. FEMS Yeast Res 4: 375–384

Hollenberg CP, Janowicz ZA (1988) DNA molecules coding for FMDH control regions and structured gene having FMDH activity and their uses. European Patent Application EPA 0.299108

Janowicz ZA, Eckart MR, Drewke C, Roggenkamp RO, Hollenberg CP, Maat J, Leenbore AM, Visser C, Verheyen CT (1985) Cloning and characterization of the *DAS* gene encoding the major methanol assimilatory enzyme from the methylotrophic yeast *Hansenula polymorpha*. Nucleic Acids Res 13: 3043–306

Janowicz SA, Meffkellbach A, Jacobs E, Harford N, Cimerbrack M, Hollenberg CP (1991) Simultaneous expression of the S and L surface antigens of hepatitis B, and formation of mixed particles in the methylotrophic yeast *Hansenula polymorpha*. Yeast 7: 431–443

Jenzelewske V (2002) Fermentation and primary product recovery. In: Gellissen G (Ed) *Hansenula polymorpha* – biology and applications. Wiley-VCH, Weinheim, pp 156–174

Jigami Y, Odani T (1999) Mannosylphosphate transfer to yeast mannan. Biochim Biophys Acta 1426: 335–345

Kang HA, Hong W-K, Sohn J-H, Choi E-S, Rhee SK (2001b) Molecular characterization of the actin-encoding gene and the use of its promoter for a dominant selection system in the methylotrophic yeast *Hansenula polymorpha*. Appl Microbiol Biotechnol 55: 734–741

Kang HA, Kang W, Hong W-K, Kim MW, Kim JY, Sohn J-H, Choi E-S, Choe K-B, Rhee SK (2001a) Development of expression systems for the production of recombinant human serum albumin using the *MOX* promoter in *Hansenula polymorpha* DL-1. Biotechnol Bioeng 76: 175–185

Kang HA, Kim J-Y, Ko S-M, Park CS, Ryu DY, Sohn J-H, Choi E-S, Rhee S-K (1998a) Cloning and characterization of the *Hansenula polymorpha* homologue of the *Saccharomyces cerevisiae PMR1* gene. Yeast 14: 1233–1240

Kang HA, Sohn J-H, Agaphonov MO, Choi E-S, Ter-Avanesyan MD, Rhee SK (2002) Development of expression systems for the production of recombinant proteins in *Hansenula polymorpha* DL-1. In: Gellissen G (Ed) *Hansenula polymorpha* – biology and applications. Wiley-VCH, Weinheim, pp 124–146

Kang HA, Sohn J-H, Choi E-S, Chung B-H, Yu M-H, Rhee S-E (1998b) Glycosylation of human α_1-antitrypsin in *Saccharomyces cerevisiae* and methylotrophic yeasts. Yeast 14: 371–381

Kiel JAKW, Hilbrands RE, van der Klei IJ, Rasmussen SW, Salomons FA, van der Heide M, Faber KN, Cregg JM, Veenhuis M (1999) *Hansenula polymorpha* Pex1p and Pex6p are peroxisome-associated AAA proteins that functionally and physically interact. Yeast 15: 1059–1078

Kiel JAKW, Keizer-Gunbink JK, Krause T, Komori M, Veenhuis M (1995) Heterologous complementation of peroxisome function in yeast, the *Saccharomyces cerevisiae PAS3* gene restores peroxisome biogenesis in a *Hansenula polymorpha per9* disruption mutant. FEBS Lett 377: 434–438

Kim MW, Agaphonov MO, Kim JY, Rhee SK,

Kang HA (2002) Sequencing and functional analysis of the *Hansenula polymorpha* genomic fragment containing the *YPT1* and *PMI40* genes. Yeast 19: 863–871

Kim MW, Rhee SK, Kim JY, Shimma YI, Chiba Y, Jigami Y, Kang HA (2004) Characterization of N-linked oligosaccharides assembled on secretory recombinant glucose oxidase and cell wall mannoproteins from the methylotrophic yeast *Hansenula polymorpha*. Glycobiology 14: 243–253

Kim SY, Sohn S-H, Kang HA, Choi E-S (2001) Cloning and characterization of the *Hansenula polymorpha* homologue of the *Saccharomyces cerevisiae MNN9* gene. Yeast 18: 455–461

Kim SY, Sohn JH, Pyun YR, Choi ES (2002) A cell surface display system using novel GPI-anchored proteins in *Hansenula polymorpha*. Yeast 19: 1153–1163

Kim SY, Sohn JH, Bae JH, Pyun YR, Agaphonov MO, Ter-Avanesyan MD, Choi ES (2003) Efficient library construction by *in vivo* recombination with a telomere-originated autonomously replicating sequence of *Hansenula polymorpha*. Appl Environ Microbiol 69: 4448–4454

Klabunde J, Dübel A, Waschkp D, Gellissen G, Hollenberg CP, Suckow M (2002) Single step co-integration of multiple expressible heterologous genes into the ribosomal DNA of the methylotrophic yeast *Hansenula polymorpha*. Appl Microbiol Biotechnol 58: 797–805

Klabunde J, Kunzi G, Gellissen G, Hollenberg CP (2003) Integration of heterologous genes in several yeast species using vectors containing a *Hansenula polymorpha*-derived rDNA targeting element. FEMS Yeast Res 4: 185–193

Knappmann S, Prets R, Gellissen G, Heiler M, Bradv GH (2000) *HARO7* encodes chorismate mutase of the methylotrophic yeast *Hansenula polymorpha* and is derepressed upon methanol utilization. J Bacteriol 182: 4188–4197

Kreger-van Rij NJW (1984) The yeasts. A taxonomic study. 3rd edition. Amsterdam. Elsevier Science Publisher, Amsterdam

Kurtzman CP, Fell JW (1998) The yeasts, a taxonomic study, 4th edition. Elsevier, Amsterdam

Kurtzman CP, Robnett CJ (1998) Identification and phylogeny of ascomycetous yeasts from analysis of nuclear large subunit (26S) ribosomal DNA partial sequences. Antonie van Leeuwenhoek 73: 331–371

Kurtzman CP (1984) Synonymy of the yeast genera *Hansenula* and *Pichia* demonstrated through comparison of deoxyribonucleic acid relatedness. Antonie van Leeuwenhoek 50: 209–217

Laptchev K (2002) Basic genetics of *Hansenula polymorpha*. In: Gellissen G (Ed) *Hansenula polymorpha* – biology and applications. Wiley-VCH, Weinheim, pp 8–20

Lederboer AM, Eveleyn L, Maat J, Visser C, Bos JW, Verrips CT, Janowicz ZA, Eckart M, Roggenkamp RO, Hollenberg CP (1985) Molecular cloning and characterization of a gene coding for methanol oxidase in *Hansenula polymorpha*. Nucleic Acids Res 13: 3063–3082

Levine DW, Cooney CL (1973) Isolation and characterization of a thermotolerant methanol-utilizing yeast. Appl Microbiol 26: 982–989

Mannazzu I, Guerra E, Strabbioli R, Masia A, Mavsvrale GR, Zonodini MA, Fatichenti F (1997) Vanadium affects vacuolation and phosphate metabolism in *Hansenula polymorpha*. FEMS Microbiol Lett. 147: 23–28

Mayer AF, Hellingwerf K, Scheffers H, Lopez-Ulibarri R, Oertel S, Dahlems U, Strasser AWM, van Loon APGM (1999) An expression system matures: a highly efficient and cost-effective process for phytase production by recombinant strains of *Hansenula polymorpha*. Biotechnol Bioeng 63: 373–381

Marri L, Ramolini GM, Satta G (1993) Chromosome polymorphism among strains of *Hansenula polymorpha* (syn. *Pichia angusta*). Appl Environ Microbiol 59: 939–943

Mewes HW, Frishman D, Güldener U, Mannhaupt G, Mayer K, Mokrejs M, Morgenstern B, Münsterkötter M, Rudd S, Weil B (2002) MIPS: a database for genomes and protein sequences. Nucleic Acids Res 30: 31–34

Middelhoven W (2002) History, habitat, variability, nomenclature and phylogenetic position of *Hansenula polymorpha*. In: Gellissen G (Ed) *Hansenula polymorpha* – biology and applications. Wiley-VCH, Weinheim, pp 1–7

Müller RG, Castellino FJ, Breithaupt BK (1997) Characterization of the acidic oligosaccharides assembled on the *Pichia pastoris*-expressed recombinant kringle 2 domain of tis-

sue-type plasminogen activator. Biotechnol Appl Biochem 26: 79–83

Montesino R, Nuñez M, Quintero O, García R, Falcón V, Cremata JA (1999) Characterization of the oligosaccharides assembled on the *Pichia pastoris*-expressed recombinant aspartic protease. Glycobiology 9: 1037–1043

Montesino R, García R, Quintero O, Cremata JA (1998) Variation in N-linked oligosaccharide structures on heterologous proteins secreted by the methylotrophic yeast *Pichia pastoris*. Protein Exp Purif 14: 197–207

Mopais JOF DE, Maia MHD (1959) Estudos de microorganismos encontrados em leitos de despejos de caldas de destilarias de Pernambuco. II. Uma nova especie de *Hansenula*: *H. polymorpha*. Anais da Escola Superior de Quimica da Universidade do Recife 1: 15–20

Müller S, Sandal T, Kamp-Hansen P, Dalboge H (1998) Comparison of expression systems in the yeasts *Saccharomyces cerevisiae*, *Hansenula polymorpha*, *Kluyveromyces lactis*, *Schizosaccharomyces pombe* and *Yarrowia lipolytica*. Cloning of two novel promoters from *Yarrowia lipolytica*. Yeast 14: 1267–1283

Müller F U, Tiebe A, Waschk D, Möhle C, Müller F J, Seigelchifer M, Pesce A, Jelezarowski V, Gellissen G (2002) Production of IFNα-2a in *Hansenula polymorpha*. Proc Biochem 38: 15–25

Ogata K, Nishikawa H, Ohsugi M (1969) Yeast capable of utilizing methanol. Agric Biol Chem 33: 1519–1520

Oh KS, Kwon O, Oh YW, Sohn MJ, Jung S, Kim YK, Kim MG, Kheir SK, Gellissen G, Kang HA (2004) Fabrication of a partial genome microarray of the methylotrophic yeast *Hansenula polymorpha*, optimization and evaluation for transcript profiling. J Microbiol Biotechnol 14. (in press)

Osoegawa K, Woon PY, Zhao B, Frengen E, Tateno M, Catanese JJ, de Jong PJ (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 52: 1–8

Osoegawa K, de Jong PJ, Frengen E, Ioannou PA (1999) Construction of bacterial artificial chromosome (BAC/PAC) libraries. Curr Protocol Hum Genet 5 15.1–5.15.33

Papendieck A, Dahlems U, Gellissen G (2002) Technical enzyme production and whole cell biocatalysis: application of *Hansenula polymorpha*. In: Gellissen G (Ed) *Hansenula polymorpha* – biology and applications. Wiley-VCH, Weinheim, pp 255–271

Park BS, Arasine V, Kim CH, Rhee SK, Kang HA (2004) Secretory production of *Zymomonas mobilis* levansucrase by the methylotropic yeast *Hansenula polymorpha*. Enzym Microbiol Tech 34: 132–138

Parpinello G, Berardi F, Strabbioli R (1998) A regulatory mutant of *Hansenula polymorpha* exhibiting methanol utilization metabolism and peroxisome proliferation in glucose. J Bacteriol 180: 2958–2967

Periyannan A, Mueckenbach A, Kelly P, Gellissen G, Hollenberg CP (1998) Cloning and characterization of the gene encoding a repressible acid phosphatase (*PHO1*) from the methylotrophic yeast *Hansenula polymorpha*. Appl Microbiol Biotechnol 50: 77–84

Rameezani-Rad M, Hollenberg CP, Lauber J, Wedler H, Gries E, Wagner C, Albermann K, Hani J, Piontek M, Dahlems U, Gellissen G (2003) The *Hansenula polymorpha* (strain CBS4732) genome sequencing and analysis. FEMS Yeast Res 4: 207–215

Raschke WC, Neidhart BR, Hendricks M, Cregg JM (1996) Inducible expression of a heterologous protein in *Hansenula polymorpha* using the alcohol oxidase 1 promoter of *Pichia pastoris*. Gene 177: 163–167

Rehm H-J, Reed G (Eds) (1993) Biotechnology. Vol 2, Genetic fundamentals and genetic engineering. VCH, Weinheim

Rodriguez L, Narciandi RE, Roca D, Cremata J, Montesinos R, Rodriguez L, Grillo JM, Muzio V, Herrera LS, Delgado JM (1996) Invertase secretion in *Hansenula polymorpha* under the *AOX1* promoter from *Pichia pastoris*. Yeast 12: 815–822

Roggenkamp R, Hansen H, Eckart M, Janowicz Z, Hollenberg CP (1986) Transformation of the methylotrophic yeast *Hansenula polymorpha* by autonomous replication and integration vectors. Mol Gen Genet 202: 302–308

Rossolini A, Romano J, Wiemken A, de Virgilio C (1999) The thermophilic yeast *Hansenula polymorpha* does not require trehalose synthesis for growth at high temperatures but does for normal acquisition of thermotolerance. J Bacteriol 181: 4665–4668

Romano I (1998) Untersuchungen zur Trehalose-6-phosphat Synthase und Klonierung und Deletion des *TPS1* Gens in der methylo-

trophen Hefe *Hansenula polymorpha*. PhD Thesis, University of Basel, Switzerland

Salomons FA, Kiel JAKW, Faber KN, Veenhuis M, van der Klei IJ (2000) Overproduction of pex3p stimulates import of alcohol oxidase and dihydroxyacetone synthase in a *Hansenula polymorpha pex14* null mutant. J Biol Chem 276: 44570–44574

Schaefer S, Piontek M, Ahn S-J, Papendieck A, Janowicz ZA, Gellissen G (2001) Recombinant hepatitis B vaccines – characterization of the viral disease and vaccine production in the methylotrophic yeast *Hansenula polymorpha*. In: Dembowsky K Stadler P (Eds), Novel therapeutic proteins: selected case studies., Wiley-VCH, Weinheim, pp 245–274

Schaefer S, Piontek M, Ahn S-J, Papendieck A, Janowicz ZA, Timmermans I, Gellissen G (2002) Recombinant hepatitis B vaccines – disease characterization and vaccine production. In: Gellissen G (Ed) *Hansenula polymorpha* – biology and applications. Wiley-VCH, Weinheim, pp 175–210

Scorer CA, Buckholz RG, Clare JJ, Romanos MA (1993) The intracellular production and secretion of HIV-1 envelop protein in the methylotrophic yeast *Pichia pastoris*. Gene 136: 111–119

Siverio JM (2002) Biochemistry and genetics of nitrate assimilation. In: Gellissen G (Ed) *Hansenula polymorpha* – biology and applications. Wiley-VCH, Weinheim, pp 21–40

Sohn J-H, Choi E-S, Kang HA, Rhee J-S, Rhee S-K (1999a) A family of telomere-associated autonomously replicating sequences and their functions in targeted recombination in *Hansenula polymorpha* DL-1. J Bacteriol 181: 1005–1013

Sohn J-H, Choi E-S, Kang HA, Rhee J-S, Agaphonov MO, Ter-Avanesyan MD, Rhee S-K (1999b) A dominant selection system designed for copy number-controlled gene integration in *Hansenula polymorpha* DL-1. Appl Microbiol Biotechnol 51: 800–807

Sohn J-H, Choi E-S, Kim C-H, Agaphonov MO, Ter-Avanesyan MD, Rhee J-S, Rhee S-K (1996) A novel autonomously replicating sequence (ARS) for multiple integration in the yeast *Hansenula polymorpha* DL-1. J Bacteriol 178: 4420–4428

Sohn J-H, Kang HA, Rao KJ, Kim CH, Choi E-S, Choung BH, Rhee S-K (2001) Current status of the anticoagulant hirudin: its biotechnological production and clinical practice. Appl Microbiol Biotechnol 57: 606–613

Spier RE (2000) Yeast as a cell factory. Enzyme Microbiol Technol 26: 639

Strahl-Bolsinger S, Gentzsch M, Tanner W (1999) Protein O-mannosylation. Biochim Biophys Acta 1426: 297–307

Suckow M, Gellissen G (2002) The expression platform based on *Hansenula polymorpha* strain RB11 and its derivatives–history, status and perspectives. In: Gellissen G (Ed) *Hansenula polymorpha* – biology and applications. Wiley-VCH, Weinheim, pp 105–123

Sudbery P (2003) Book review: Gellissen G (Ed) *Hansenula polymorpha* – biology and Application. Yeast 20: 107–1308

Tan X, Titorenko VI, van der Klei IJ, Sulter GJ, Harder W, Veenhuis M, Cregg JM (1995) Characterization of peroxisome-deficient mutants of *Hansenula polymorpha*. Curr Genet 28: 248–257

Tani Y (1984) Microbiology and biochemistry of the methylotrophic yeasts. In: Hou CT (ed) Methylotrophic microbiology, biochemistry, and genetics. CRC Press, Boca Raton, FL, pp 55–86

Thorisson DI, Hall HH, Wickerham LJ (1960) *Hansenula augusta*, an excellent species for demonstration of the coexistence of haploid and diploid cells in a homothallic yeast. Mycologia 52: 184–188

Thömmes J, Halbar M, Gieren H, Curvers S, Takors R, Brunschier R, Kula M-R (2001) Human chymotrypsinogen B production from *Pichia pastoris* by integrated development of fermentation and downstream processing Part 2. Protein recovery. Biotechnol Progr 17: 503–512

Tikhomirova LP, Ironomova RN, Kuznetsova EN, Fodor U, Bystrykh LV, Aminova LR, Trotsenko YA (1988) Transformation of methylotrophic yeast *Hansenula polymorpha*: cloning and expression of genes. J Basic Microbiol 28: 343–351

van Dijk R, Faber KN, Kiel JAKW, Veenhuis M, van der Klei IJ (2000) The methylotrophic yeast *Hansenula polymorpha*: a versatile cell factory. Enzyme Microbiol Technol 26: 793–800

van Dijk R, Faber KN, Hammond AT, Glick BS, Veenhuis M, Kiel JAKW (2001) Tagging *Hansenula polymorpha* genes by ran-

dom integration of linear DNA fragments (RALF). Mol Genet Genomics 266: 646–656

VAN DER KLEI IJ, VEENHUIS M (2002) *Hansenula polymorpha*: a versatile model organism in peroxisome research. In: Gellissen G (Ed) *Hansenula polymorpha* – biology and applications. Wiley-VCH, Weinheim, pp 76–94

VERAS RA, GUTIERREZ ML, VAN EIJK HM, SUNDERDY PB, VERRIPS CT (1992) Development of a strain of *Hansenula polymorpha* for the efficient expression of guar alpha-galactosidase. Yeast 8: 361–372

WASCHER D, KLABUNDE J, SUCKOW M, HOLLENBERG CP (2002) Characteristics of the *Hansenula polymorpha* genome. In: Gellissen G (Ed) *Hansenula polymorpha* – biology and applications. Wiley-VCH, Weinheim, pp 95–104

WEYDEMANN U, KEUP P, PIONTEK M, STRASSER AWM, SCHWEIGER J, GELLISSEN G, JANOWICZ SA (1995) High-level secretion of hirudin by *Hansenula polymorpha*-authentic processing of three different preprohirudins. Appl Microbiol Biotechnol 44: 377–385

WICKERHAM LJ (1951) Taxonomy of yeasts. Technical bulletin No 1029, US Dept Agric, Washington DC, pp 1–56

YURIMOTO H, SAKAI Y, KATO N (2002) Methanol metabolism. In: Gellissen G (Ed) *Hansenula polymorpha* – biology and applications. Wiley-VCH, Weinheim, pp 61–75

ZORER C, KUBIS E, KEUP P, HÖRLEIN D, BRONNE J, THÖMMES J, KULA M-R, HOLLENBERG CP, GELLISSEN G (1996) Production of two aprotinin variants in *Hansenula polymorpha*. Process Biochem 31: 679–689